

OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/836,470

DATE: 11/20/2001
 TIME: 19:11:34

Input Set : A:\836470.txt
 Output Set: N:\CRF3\11202001\I836470.raw

ENTERED

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Witte, Owen N.
 6 Weng, Zhigang
 8 (ii) TITLE OF INVENTION: IDENTIFICATION OF A G PROTEIN-COUPLED
 9 RECEPTOR TRANSCRIPTIONALLY REGULATED BY PROTEIN
 10 TYROSINE KINASE SIGNALING IN HEMATOPOIETIC CELLS
 12 (iii) NUMBER OF SEQUENCES: 40
 14 (iv) CORRESPONDENCE ADDRESS:
 15 (A) ADDRESSEE: Knobbe, Martens, Olson & Bear
 16 (B) STREET: 620 Newport Center Drive, 16th Floor
 17 (C) CITY: Newport Beach
 18 (D) STATE: CA
 19 (E) COUNTRY: U.S.A.
 20 (F) ZIP: 92660
 22 (v) COMPUTER READABLE FORM:
 23 (A) MEDIUM TYPE: Diskette
 24 (B) COMPUTER: IBM Compatible
 25 (C) OPERATING SYSTEM: DOS
 26 (D) SOFTWARE: FastSEQ for Windows Version 2.0
 28 (vi) CURRENT APPLICATION DATA:
 C--> 29 (A) APPLICATION NUMBER: US/09/836,470
 C--> 30 (B) FILING DATE: 18-Apr-2001
 31 (C) CLASSIFICATION:
 33 (vii) PRIOR APPLICATION DATA:
 34 (A) APPLICATION NUMBER: US/08/969,815
 35 (B) FILING DATE:
 37 (viii) ATTORNEY/AGENT INFORMATION:
 38 (A) NAME: Bartfeld, Neil S
 39 (B) REGISTRATION NUMBER: 39,901
 40 (C) REFERENCE/DOCKET NUMBER: UCLA015.001A
 42 (ix) TELECOMMUNICATION INFORMATION:
 43 (A) TELEPHONE: 619-235-8550
 44 (B) TELEFAX: 619-235-0176
 45 (C) TELEX:
 48 (2) INFORMATION FOR SEQ ID NO: 1:
 50 (i) SEQUENCE CHARACTERISTICS:
 51 (A) LENGTH: 1507 base pairs
 52 (B) TYPE: nucleic acid
 53 (C) STRANDEDNESS: single
 54 (D) TOPOLOGY: linear
 56 (ii) MOLECULE TYPE: cDNA
 57 (ix) FEATURE:
 59 (A) NAME/KEY: Coding Sequence
 60 (B) LOCATION: 147...1292
 61 (D) OTHER INFORMATION:
 63 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,470

DATE: 11/20/2001

TIME: 19:11:34

Input Set : A:\836470.txt

Output Set: N:\CRF3\11202001\I836470.raw

65	AAACCTCCCA GCTGGGCCTG CAGAGGGGTG CTCAGCCCTG CCTCAGGACG GGCCTGCCCT	60
66	GTGCTGCCTC AGGACTGGCT TGGGTCATTT TAAGCTGCCA GAGCCACCTT CACAAGGGGG	120
67	TCCACAGAAC TCACATAGGA GCCACC ATG AGA TCA GAA CCT ACC AAT GCA GCA	173
68	Met Arg Ser Glu Pro Thr Asn Ala Ala	
69	1 5	
71	GGA AAC ACC ACA CTG GGG GTT ACC TCC GTT CTT CAG AGC ACC TCA GTA	221
72	Gly Asn Thr Thr Leu Gly Val Thr Ser Val Leu Gln Ser Thr Ser Val	
73	10 15 20 25	
75	CCT TCT TCT GAG ACC TGC CAC GTC TCC TAC GAG GAG AGC AGA GTG GTC	269
76	Pro Ser Ser Glu Thr Cys His Val Ser Tyr Glu Glu Ser Arg Val Val	
77	30 35 40	
79	CTG GTG GTG GTG TAC AGT GCC GTG TGC CTG CTG GGC CTA CCA GCC AAC	317
80	Leu Val Val Val Tyr Ser Ala Val Cys Leu Leu Gly Leu Pro Ala Asn	
81	45 50 55	
83	TGC CTG ACT GCC TGG CTG ACG CTG CTG CAA GTC CTG CAG AGG AAC GTG	365
84	Cys Leu Thr Ala Trp Leu Thr Leu Leu Gln Val Leu Gln Arg Asn Val	
85	60 65 70	
87	CTA GCC GTC TAC CTG TTC TGC CTG TCC CTC TGT GAG CTG CTC TAC ATC	413
88	Leu Ala Val Tyr Leu Phe Cys Leu Ser Leu Cys Glu Leu Leu Tyr Ile	
89	75 80 85	
91	AGC ACG GTG CCA TTG TGG ATC ATC TAC ATC CAG AAT CAG CAC AAA TGG	461
92	Ser Thr Val Pro Leu Trp Ile Ile Tyr Ile Gln Asn Gln His Lys Trp	
93	90 95 100 105	
95	AAC CTG GGT CCG CAG GCC TGC AAG GTG ACT GCT TAC ATC TTC TTC TGC	509
96	Asn Leu Gly Pro Gln Ala Cys Lys Val Thr Ala Tyr Ile Phe Phe Cys	
97	110 115 120	
99	AAC ATC TAC ATC AGC ATC CTC TTG CTC TGC TGC ATT TCC TGC GAC CGC	557
100	Asn Ile Tyr Ile Ser Ile Leu Leu Leu Cys Cys Ile Ser Cys Asp Arg	
101	125 130 135	
103	TAC ATG GCC GTG GTC TAT GCA CTG GAG AGC CGA GGC CAC CGC CAC CAG	605
104	Tyr Met Ala Val Val Tyr Ala Leu Glu Ser Arg Gly His Arg His Gln	
105	140 145 150	
107	AGG ACT GCT GTC ACC ATT TCT GCG TGT GTG ATT CTT CTT GTT GGA CTT	653
108	Arg Thr Ala Val Thr Ile Ser Ala Cys Val Ile Leu Leu Val Gly Leu	
109	155 160 165	
111	GTT AAC TAT CCA GTG TTT GAC ATG AAG GTG GAG AAG AGT TTC TGC TTT	701
112	Val Asn Tyr Pro Val Phe Asp Met Lys Val Glu Lys Ser Phe Cys Phe	
113	170 175 180 185	
115	GAG CCC CTG AGG ATG AAC AGC AAG ATA GCC GGC TAC CAC TAC CTG CGT	749
116	Glu Pro Leu Arg Met Asn Ser Lys Ile Ala Gly Tyr His Tyr Leu Arg	
117	190 195 200	
119	TTC ACC TTT GGC TTT GCC ATC CCT CTC GGC ATC CTG GCG TTC ACC AAT	797
120	Phe Thr Phe Gly Phe Ala Ile Pro Leu Gly Ile Leu Ala Phe Thr Asn	
121	205 210 215	
123	CAC CAG ATC TTC CGG AGC ATC AAA CTC AGT GAC AGC CTG AGC GCT GCG	845
124	His Gln Ile Phe Arg Ser Ile Lys Leu Ser Asp Ser Leu Ser Ala Ala	
125	220 225 230	
127	CAG AAG AAC AAG GTG AAG CGC TCC GCC ATC GCG GTC ACC ATC TTC	893
128	Gln Lys Asn Lys Val Lys Arg Ser Ala Ile Ala Val Val Thr Ile Phe	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,470

DATE: 11/20/2001

TIME: 19:11:34

Input Set : A:\836470.txt

Output Set: N:\CRF3\11202001\I836470.raw

```

129      235      240      245
131 CTG GTC TGC TTT GCT CCC TAC CAC GTG GTA CTC CTC GTC AAA GCT GCC 941
132 Leu Val Cys Phe Ala Pro Tyr His Val Val Leu Leu Val Lys Ala Ala
133 250      255      260      265
135 AGC TTT TCC TTC TAC CAA GGA GAC ATG GAT GCC GTG TGT GCC TTT GAA 989
136 Ser Phe Ser Phe Tyr Gln Gly Asp Met Asp Ala Val Cys Ala Phe Glu
137      270      275      280
139 AGC AGA CTG TAC ACA GTC TCT ATG GTG TTT CTG TGC CTG TCT ACA GTC 1037
140 Ser Arg Leu Tyr Thr Val Ser Met Val Phe Leu Cys Leu Ser Thr Val
141      285      290      295
143 AAC AGT GTG GCT GAC CCC ATC ATC TAC GTG CTG GGT ACA GAC CAC TCT 1085
144 Asn Ser Val Ala Asp Pro Ile Ile Tyr Val Leu Gly Thr Asp His Ser
145      300      305      310
147 CGG CAA GAA GTG TCC AGA ATC CAC ACA GGG TGG AAA AAG TGG TCC ACA 1133
148 Arg Gln Glu Val Ser Arg Ile His Thr Gly Trp Lys Lys Trp Ser Thr
149      315      320      325
151 AAG ACA TAT GTT ACA TGC TCA AAG GAC TCT GAG GAG ACA CAC TTG CCC 1181
152 Lys Thr Tyr Val Thr Cys Ser Lys Asp Ser Glu Glu Thr His Leu Pro
153 330      335      340      345
155 ACA GAG CTT TCA AAC ACA TAC ACC TTC CCC AAT CCC GCG CAC CCT CCA 1229
156 Thr Glu Leu Ser Asn Thr Tyr Thr Phe Pro Asn Pro Ala His Pro Pro
157      350      355      360
159 GGA TCA CAG CCA GCG AAG CTA GGT TTA CTG TGC TCG CCA GAG AGA CTG 1277
160 Gly Ser Gln Pro Ala Lys Leu Gly Leu Leu Cys Ser Pro Glu Arg Leu
161      365      370      375
163 CCT GAG GAG CTC TGC TAAGAGACGA TTGTCCACTC TTCCTCAAAA CTAGCACCAG T 1333
164 Pro Glu Glu Leu Cys
165      380
167 CACACATACC TGGTCCTCTG AGTCACCGTC TGGGGTGTCC ACAGCACTAT AGATGCCTTT 1393
168 GTTCGGGCAC ACGCTGCTGA TCTTTCCTTC CTAAGGCCAC CAACTCTGAA AGTATCTGTT 1453
169 CCTTAAACTG TCCTCAGGCT CCCCTCTATG GAAAGCGGGG CTTGCTAAGG GACC 1507
171 (2) INFORMATION FOR SEQ ID NO: 2:
173 (i) SEQUENCE CHARACTERISTICS:
174 (A) LENGTH: 382 amino acids
175 (B) TYPE: amino acid
176 (C) STRANDEDNESS: single
177 (D) TOPOLOGY: linear
179 (ii) MOLECULE TYPE: protein
180 (v) FRAGMENT TYPE: internal
182 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
184 Met Arg Ser Glu Pro Thr Asn Ala Ala Gly Asn Thr Thr Leu Gly Val
185 1 5 10 15
186 Thr Ser Val Leu Gln Ser Thr Ser Val Pro Ser Ser Glu Thr Cys His
187 20 25 30
188 Val Ser Tyr Glu Glu Ser Arg Val Val Leu Val Val Val Tyr Ser Ala
189 35 40 45
190 Val Cys Leu Leu Gly Leu Pro Ala Asn Cys Leu Thr Ala Trp Leu Thr
191 50 55 60
192 Leu Leu Gln Val Leu Gln Arg Asn Val Leu Ala Val Tyr Leu Phe Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,470

DATE: 11/20/2001

TIME: 19:11:34

Input Set : A:\836470.txt

Output Set : N:\CRF3\11202001\I836470.raw

```

193 65          70          75          80
194 Leu Ser Leu Cys Glu Leu Leu Tyr Ile Ser Thr Val Pro Leu Trp Ile
195          85          90          95
196 Ile Tyr Ile Gln Asn Gln His Lys Trp Asn Leu Gly Pro Gln Ala Cys
197          100          105          110
198 Lys Val Thr Ala Tyr Ile Phe Phe Cys Asn Ile Tyr Ile Ser Ile Leu
199          115          120          125
200 Leu Leu Cys Cys Ile Ser Cys Asp Arg Tyr Met Ala Val Val Tyr Ala
201          130          135          140
202 Leu Glu Ser Arg Gly His Arg His Gln Arg Thr Ala Val Thr Ile Ser
203          145          150          155          160
204 Ala Cys Val Ile Leu Leu Val Gly Leu Val Asn Tyr Pro Val Phe Asp
205          165          170          175
206 Met Lys Val Glu Lys Ser Phe Cys Phe Glu Pro Leu Arg Met Asn Ser
207          180          185          190
208 Lys Ile Ala Gly Tyr His Tyr Leu Arg Phe Thr Phe Gly Phe Ala Ile
209          195          200          205
210 Pro Leu Gly Ile Leu Ala Phe Thr Asn His Gln Ile Phe Arg Ser Ile
211          210          215          220
212 Lys Leu Ser Asp Ser Leu Ser Ala Ala Gln Lys Asn Lys Val Lys Arg
213          225          230          235          240
214 Ser Ala Ile Ala Val Val Thr Ile Phe Leu Val Cys Phe Ala Pro Tyr
215          245          250          255
216 His Val Val Leu Leu Val Lys Ala Ala Ser Phe Ser Phe Tyr Gln Gly
217          260          265          270
218 Asp Met Asp Ala Val Cys Ala Phe Glu Ser Arg Leu Tyr Thr Val Ser
219          275          280          285
220 Met Val Phe Leu Cys Leu Ser Thr Val Asn Ser Val Ala Asp Pro Ile
221          290          295          300
222 Ile Tyr Val Leu Gly Thr Asp His Ser Arg Gln Glu Val Ser Arg Ile
223          305          310          315          320
224 His Thr Gly Trp Lys Lys Trp Ser Thr Lys Thr Tyr Val Thr Cys Ser
225          325          330          335
226 Lys Asp Ser Glu Glu Thr His Leu Pro Thr Glu Leu Ser Asn Thr Tyr
227          340          345          350
228 Thr Phe Pro Asn Pro Ala His Pro Pro Gly Ser Gln Pro Ala Lys Leu
229          355          360          365
230 Gly Leu Leu Cys Ser Pro Glu Arg Leu Pro Glu Glu Leu Cys
231          370          375          380

```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2938 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 901...2040

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,470

DATE: 11/20/2001

TIME: 19:11:34

Input Set : A:\836470.txt

Output Set: N:\CRF3\11202001\I836470.raw

246 (D) OTHER INFORMATION:

248 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

250	GGGAGGGGTG	CNANGCTAGC	CACGCAGGCG	GGGCCCTGGG	TCATTTTAAN	CTCTCAGAGT	60
251	GAACGTCTTG	ATAGGACCGA	CAANACNCAT	NACNTGTACT	TAGATAGCTT	ATCTTANANC	120
252	CACNCTGANA	TTGGAACCCG	CAAAATATGC	CNGGGAGGAA	GGTGAGCAAG	GGACACGACA	180
253	CTCACCCGGA	TAAACCCAAC	AAGCGCAGCG	AGGCTGTGGG	GAAACCGGAN	CCCTGCACAC	240
254	CGCCGGGGGA	AGGTGGGCCN	CCGCCACCAC	CGTGGAAGAA	CAGCGCGGAN	GCACCCACG	300
255	AGATGAGACG	GAAGTCCCGT	GAGATCCAGC	AATNCCNACT	GTGGGTCTGA	CCCAGGATAN	360
256	CGGAAAGCAG	GGACGTGAAC	AGCCCTCCTC	ATGTTCTTGA	CACCGTCATT	CTCAGCAGCT	420
257	CAGCTAAGGC	ACAGAGGCAG	CCGAGCGTCT	GTCAGCAGAG	TCGTGGCTGA	GCAGAACACG	480
258	CCACACGCCA	CACGCCACAC	GCCACACGTG	CAGGATTGCT	CAAGATGGAA	GGGCACAGTG	540
259	GAATATATAT	ATATATTTAT	ATTTTGGCG	AGACCCTGGA	GGACACACTG	AATACAATGG	600
260	AATACCATCC	CGCCTTTGAA	AGGAAGGGAA	ATCCTGGCAC	ACGCTGCAAC	AGGAGGGAGC	660
261	TTGAGGACAC	TGTGGTGAGT	GGAGCACGTG	AGACACGGAA	GGACACACGC	TGAAGACACG	720
262	CAGAGATGCC	CACCCACGTG	GGGAGGTGAC	AGGGGAGCCC	AGCGCACAGA	GACAAAGTGG	780
263	AATGGAGGCC	TGGGGGCTGG	GAGCAAATGC	GGAGCGAGTG	CTTCCTGGGG	CAGAGTCTCC	840
264	GTTTGGGAAG	ATGAGAAGGT	TCTGCCGACG	GATGCTGGCG	ATGGTTGCAG	AAGAATGTGA	900
265	ATG TGC CCA	ATG CTA CTG	AAA AAC	GGT TAC AAT	GGA AAC GCC	ACC CCA	948
266	Met Cys Pro	Met Leu Leu	Lys Asn	Gly Tyr Asn	Gly Asn Ala	Thr Pro	
267	1	5	10	15			
269	GTG ACC ACC	ACT GCC CCG	TGG GCC TCC	CTG GGC CTC	TCC GCC AAG	ACC	996
270	Val Thr Thr	Thr Ala Pro	Trp Ala Ser	Leu Gly Leu	Ser Ala Lys	Thr	
271		20	25	30			
273	TGC AAC AAC	GTG TCC TTC	GAA GAG AGC	AGG ATA GTC	CTG GTC GTG	GTG	1044
274	Cys Asn Asn	Val Ser Phe	Glu Glu Ser	Arg Ile Val	Leu Val Val	Val	
275		35	40	45			
277	TAC AGC GCG	GTG TGC ACG	CTG GGG GTG	CCG GCC AAC	TGC CTG ACT	GCG	1092
278	Tyr Ser Ala	Val Cys Thr	Leu Gly Val	Pro Ala Asn	Cys Leu Thr	Ala	
279		50	55	60			
281	TGG CTG GCG	CTG CTG CAG	GTA CTG CAG	GGC AAC GTG	CTG GCC GTC	TAC	1140
282	Trp Leu Ala	Leu Leu Gln	Val Leu Gln	Gly Asn Val	Leu Ala Val	Tyr	
283		65	70	75		80	
285	CTG CTC TGC	CTG GCA CTC	TGC GAG CTG	CTG TAC ACA	GGC ACG CTG	CCA	1188
286	Leu Leu Cys	Leu Ala Leu	Cys Glu Leu	Leu Tyr Thr	Gly Thr Leu	Pro	
287		85	90	95			
289	CTC TGG GTC	ATC TAT ATC	CGC AAC CAG	CAC CGC TGG	ACC CTA GGC	CTG	1236
290	Leu Trp Val	Ile Tyr Ile	Arg Asn Gln	His Arg Trp	Thr Leu Gly	Leu	
291		100	105	110			
293	CTG GCC TGC	AAG GTG ACC	GCC TAC ATC	TTC TTC TGC	AAC ATC TAC	GTC	1284
294	Leu Ala Cys	Lys Val Thr	Ala Tyr Ile	Phe Phe Cys	Asn Ile Tyr	Val	
295		115	120	125			
297	AGC ATC CTC	TTC CTG TGC	TGC ATC TCC	TGC GAC CGC	TTC GTG GCC	GTG	1332
298	Ser Ile Leu	Phe Leu Cys	Cys Ile Ser	Cys Asp Arg	Phe Val Ala	Val	
299		130	135	140			
301	GTG TAC GCG	CTG GAG AGT	CGG GGC CGC	CGC CGG AGG	ACC GCC ATC		1380
302	Val Tyr Ala	Leu Glu Ser	Arg Gly Arg	Arg Arg Arg	Arg Thr Ala	Ile	
303		145	150	155		160	
305	CTC ATC TCC	GCC TGC ATC	TTC ATC CTC	GTC GGG ATC	GTT CAC TAC	CCG	1428
306	Leu Ile Ser	Ala Cys Ile	Phe Ile Leu	Val Gly Ile	Val His Tyr	Pro	

VERIFICATION SUMMARY

DATE: 11/20/2001

PATENT APPLICATION: US/09/836,470

TIME: 19:11:35

Input Set : A:\836470.txt

Output Set: N:\CRF3\11202001\I836470.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]